sustainability

MDPI

Article

Workplace Flexibility, Health Status, and Subjective Work Performance: Evidence from Japanese Panel Data

Shotaro Watanabe

College of International Management, Ritsumeikan Asia Pacific University, 1-1 Jumonjibaru, Beppu 874-8577, Oita, Japan; s-wata@apu.ac.jp

Abstract

Workplace flexibility may have different implications for work performance depending on its form and workers' circumstances. Using nationally representative panel data from the Japan Household Panel Survey and Keio Household Panel Survey (JHPS/KHPS), this study examines whether schedule- and location-based flexibility are associated with subjective work performance and whether these associations vary by health status. The baseline analysis uses 7678 person-year observations from 3259 workers aged 20–64 in 2019–2021 and estimates random-effects ordered logit models. Poor self-rated health is consistently associated with lower subjective work performance. Schedule-Based Flexibility (SBF) shows the clearest positive baseline pattern: it is associated with a 10.2 percentage-point increase in the probability of high subjective work performance (SWP ≥ 7) among workers reporting poor health, compared with 2.5 percentage points among those not reporting poor health; the difference in average marginal effects is 7.7 percentage points ($p = 0.035$). However, the SBF × poor-health interaction turns negative in 2022–2024 and is statistically insignificant in the pooled 2019–2024 analysis, indicating substantial temporal heterogeneity. The remote-work estimates are statistically imprecise. The findings suggest SBF may be relevant to sustainable, health-supportive work design, but warrant cautious, noncausal interpretation.

Keywords: workplace flexibility; schedule-based flexibility; remote work; subjective work performance; self-rated health; work design; human sustainability; sustainable work design; Japan

1. Introduction

Japan's labor market is facing persistent labor-supply constraints. Projections by the National Institute of Population and Social Security Research indicate that the working-age population will decline from approximately 75 million in 2020 to 62 million by 2040 [1], while the Ministry of Health, Labour and Welfare characterizes labor shortages as a long-term challenge [2]. Against this background, maintaining organizational functioning requires attention not only to the recruitment of additional workers but also to the design of work arrangements that can accommodate heterogeneous worker needs and constraints. In Japan, longitudinal evidence links long working hours to poorer mental health, underscoring the importance of work organization for employee health [3]. Survey evidence further indicates that more Japanese workers perceived disadvantages than advantages in disclosing an illness at work, highlighting potential barriers to communicating health-related needs in the workplace [4]. Workplace flexibility is relevant in this context because it may provide employees with greater discretion over how work is organized while also creating new

Check for updates

Academic Editors: Kirk Chang and Susan Akinwalere

Received: 29 September 2026
Revised: 24 November 2026
Accepted: 24 November 2026
Published: 27 November 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

demands. Understanding which forms of flexibility are associated with better work outcomes, and for whom, is therefore relevant to both employee well-being and the design of sustainable work systems.

Japan's experience with telework also illustrates why workplace flexibility should not be reduced to a simple choice between remote and office-based work. Telework expanded rapidly during the COVID-19 pandemic, but its use subsequently declined before showing signs of stabilization. According to the latest national survey, 16.8% of employed workers had engaged in telework during the preceding year in fiscal 2025, while substantial regional differences in telework adoption have persisted [5]. These patterns suggest that location-based flexibility has become neither universal nor uniformly embedded in Japanese workplaces. More importantly, a post-pandemic debate centered primarily on remote versus office work may overlook other dimensions of flexibility, particularly workers' discretion over when work is performed. Evaluating flexible work therefore requires distinguishing flexibility over work location from flexibility over working time and asking under what conditions each is associated with workers' ability to perform their jobs.

The implications of workplace flexibility may depend both on the form of flexibility and on the constraints workers face. Recent research emphasizes that temporal and spatial flexibility, as well as formal availability and actual use, should not be treated as equivalent forms of flexible work [6,7]. From a Job Demands–Resources perspective, flexible work arrangements can provide resources by increasing workers' discretion over when or where work is performed, but they can also introduce demands related to self-management, coordination, and work–nonwork boundaries [8,9]. These trade-offs may be particularly important for workers in poorer health. Prior evidence suggests that health can condition workers' responses to flexible work arrangements, while recent research on workers with health conditions also points to heterogeneous work-related consequences of telework [10,11]. Yet less is known about whether health status conditions the association between distinct temporal and spatial forms of workplace flexibility and work performance.

This study addresses this gap using nationally representative panel data from the Japan Household Panel Survey and the Keio Household Panel Survey (JHPS/KHPS), with 2019–2021 serving as the baseline observation period. It examines whether distinct forms of workplace flexibility are associated with subjective work performance and whether these associations vary by workers' health status. The analysis distinguishes Schedule-Based Flexibility (SBF), capturing reported working-time arrangements that permit flexibility or discretion over working hours, from location-based flexibility. For remote work, it further distinguishes policy-only Remote-Work Availability (RWA) from reported Remote-Work Usage (RWU), while also considering Remote Work Exposure (RWE) as a broader benchmark. This design allows the study to compare flexibility over when and where work is performed within a common empirical framework and to assess whether these associations differ by health status.

This study makes two main contributions. First, it extends research on workplace flexibility by examining workers' health status as a boundary condition for the relationship between distinct forms of flexibility and subjective work performance. Rather than treating flexibility as a single practice, the analysis compares flexibility over working time with flexibility over work location and distinguishes formal remote-work availability from reported usage. The contribution therefore lies not in establishing the multidimensionality of flexibility itself, but in examining whether these dimensions show different patterns of association with work performance across health-status groups. Second, the study provides panel-based evidence from Japan, where post-pandemic debates over workplace flexibility are unfolding against persistent labor-supply constraints. Supplementary analyses also reveal substantial temporal variation in the health-related SBF association, qualifying the

generalizability of the baseline findings. By linking heterogeneous worker needs to different forms of work design, the study speaks to the human and social dimensions of sustainability and to the design of workplaces that can accommodate employee health constraints while maintaining organizational functioning.

2. Theory and Literature Review

2.1. Workplace Flexibility as a Work Design Resource and Demand

Workplace flexibility can be understood as a dimension of work design that alters employees' discretion over when and where work is performed. Work design research emphasizes that job autonomy and the temporal and physical organization of work can shape employees' motivation, strain, and performance [12,13]. For the present study, the distinction between temporal and spatial flexibility is central. Schedule-based flexibility concerns discretion over the timing of work, whereas remote-work arrangements primarily concern the location from which work is performed. These dimensions of flexibility need not provide the same resources or impose the same demands on workers.

Prior research suggests that discretion over working time can be relevant to employee outcomes. Meta-analytic evidence on flexible and compressed work schedules reports generally favorable associations with outcomes such as performance and absenteeism, although the magnitude of these associations varies across arrangements and organizational settings [14]. A systematic review of worktime control likewise finds associations with work–nonwork balance, health and well-being, and job-related outcomes, while emphasizing that causal evidence remains limited [15]. Taken together, this literature supports treating temporal discretion as a potentially important work-design resource, without assuming that its consequences are uniformly positive or causal.

The Job Demands–Resources (JD–R) model provides a useful framework for interpreting these relationships. The model distinguishes job demands that require sustained effort from job resources that help employees achieve work goals, manage demands, or maintain motivation [8,9]. From this perspective, workplace flexibility may function as a resource when it gives workers greater discretion over work timing or location, helping them manage scheduling, commuting, or personal constraints. At the same time, flexibility may generate additional demands through self-management, blurred work–nonwork boundaries, availability expectations, or coordination difficulties. Its implications therefore depend on the particular form of flexibility and on the circumstances under which workers use or experience it.

This work-design perspective also connects to the concept of sustainable work systems, which emphasizes preserving workers' health, capabilities, motivation, and participation alongside organizational functioning [16]. Workplace flexibility may therefore be relevant not only as a short-term work arrangement but also as a potential design resource for supporting workers under heterogeneous constraints. Whether this relevance differs by workers' health status is central to the present study.

2.2. Health Status as a Boundary Condition

Poor health can constrain work functioning even among individuals who remain employed and actively working. Research in labor economics documents the relationship between health and labor-market capacity [17], while the literature on presenteeism shows that health limitations can reduce on-the-job functioning without necessarily resulting in absence from work [18]. Such constraints may reduce physical or cognitive capacity and increase fatigue or recovery needs, making ordinary job demands more difficult to manage. In this study, poor health refers to respondents' general self-rated health status and is conceptually distinct from disability status. Because this measure does not establish

whether respondents have a disability or require workplace accommodations, the findings should not be interpreted as disability-specific evidence.

Subjective assessments of work performance have also been used to capture health-related variation in workers' functioning. The World Health Organization Health and Work Performance Questionnaire (HPQ), for example, uses self-reported work performance to assess productivity loss associated with health problems and has shown correspondence with archival measures of work performance across several occupational groups [19]. Such measures are therefore useful for examining variation in workers' perceived performance that may not be fully captured by absence from work alone.

Health-related constraints may make the demands associated with commuting and fixed work schedules particularly salient. Long commutes have been associated with adverse health and well-being outcomes [20,21], while limited control over work timing may restrict workers' ability to adjust work rhythms around fatigue or recovery needs. These considerations suggest that temporal and spatial flexibility may be relevant through different channels: schedule-based flexibility changes the scope for adjusting when work is performed, whereas remote work changes where work is performed and may reduce commuting demands.

Existing research provides direct evidence that health can condition workers' responses to flexible work arrangements. Rudolph and Baltes [10], for example, show across two studies that the relationship between flexible work arrangements and work engagement varies jointly with workers' age and health, using both functional-health and health-symptom measures. Recent evidence focusing specifically on workers with health conditions likewise suggests that telework may have heterogeneous associations with absenteeism, presenteeism, and return to work, with particularly mixed evidence for presenteeism [11].

Taken together, these studies suggest that health may shape workers' responses to flexible work arrangements, but an important gap remains. Less is known about whether workers' health status conditions the association between distinct forms of workplace flexibility and work performance, particularly when temporal and spatial flexibility—and, within remote work, policy availability and reported usage—are distinguished. From a work-design perspective, health status can therefore be considered a boundary condition in assessing how different forms of flexibility relate to subjective work performance. The present study examines this health-related heterogeneity directly.

2.3. Distinguishing Time-Based and Location-Based Workplace Flexibility

Empirical evidence on the consequences of remote and flexible work arrangements remains mixed and context-dependent. Meta-analytic and experimental evidence suggests that telecommuting or working from home can be associated with favorable employee and performance outcomes in some settings [22,23], while more recent randomized evidence shows that hybrid work can improve retention without reducing measured performance [24]. At the same time, remote work can create coordination challenges by reshaping communication and collaboration [25]. Taken together, this evidence suggests that the implications of workplace flexibility depend on its form and context, motivating a distinction between temporal and spatial flexibility.

One implication is the need to distinguish temporal from spatial flexibility. Flextime concerns discretion over when work is performed, whereas flexplace concerns where work is performed. Harrop et al. [7], in a meta-analysis of 113 studies, show that associations with employee outcomes differ across flextime, flexplace, and arrangements combining both; combined arrangements generally show more favorable outcomes, and flextime compares more favorably with flexplace across several outcomes. For task performance specifically, the clearest positive association was observed for arrangements combining flextime and

flexplace. These findings support analyzing temporal and spatial flexibility separately rather than treating workplace flexibility as a single construct.

A second distinction concerns the formal availability of flexible work arrangements and their actual use. Formal availability does not necessarily mean that workers can or will use an arrangement in practice. Kossek et al. [6] conceptualize flexible-work implementation as a process involving policy availability, access experiences, actual use, and outcomes, emphasizing that supervisors, organizational culture, job demands, and other contextual factors can shape whether formally available flexibility becomes usable in practice. Harrop et al. [7] likewise show that associations with employee outcomes differ between perceived availability and actual use, supporting the need to distinguish policy-only availability from reported usage.

Remote-work access and usage are also highly selective. During the early COVID-19 period, telework adoption expanded unevenly across workers and job types, while the feasibility of working from home differed substantially by occupation and task content [26,27]. Observed remote-work usage therefore reflects not only workers' preferences but also job characteristics, organizational practices, managerial decisions, and other constraints, complicating direct comparisons between users and non-users.

Evidence from Japan also illustrates the context-dependent and evolving implications of remote work. Morikawa [28] reports lower productivity under working from home than at the usual workplace during the early COVID-19 period, while subsequent studies document improvements over time, consistent with learning, adaptation, and selective continuation of remote work [29,30]. These findings reinforce the view that the implications of remote work depend on how arrangements are implemented and adapted over time rather than on remote-work status alone.

Taken together, prior research indicates that workplace flexibility is multidimensional: temporal and spatial flexibility, as well as formal availability and actual use, need not have equivalent implications. Less is known about whether the associations between these distinct forms of flexibility and subjective work performance vary by workers' health status. This gap motivates the hypotheses and exploratory comparisons developed in the next section.

3. Hypotheses

Building on the work-design and health-contingency perspectives developed above, this section develops hypotheses concerning the associations of health status and schedule-based flexibility with subjective work performance. Subjective work performance is particularly relevant in this context because it captures workers' own assessments of their recent work performance rather than relying solely on externally observed output measures.

Poor health is expected to be negatively associated with subjective work performance because health-related limitations can reduce workers' capacity to meet ordinary job demands. Consistent with the evidence reviewed in Section 2.2, this leads to the first hypothesis:

H1. *Poor health is negatively associated with subjective work performance.*

Schedule-based flexibility is expected to be positively associated with subjective work performance because working-time arrangements that provide greater scope for discretion over work timing may function as a job resource, allowing workers to better align work rhythms with personal constraints. Prior research on flexible scheduling and worktime control provides support for favorable job-related associations, while also indicating that these relationships are not uniformly positive or necessarily causal. This leads to the second hypothesis:

H2. *Schedule-based flexibility is positively associated with subjective work performance.*

The positive association between schedule-based flexibility and subjective work performance is expected to be stronger among workers reporting poor health. Health-related constraints may make fixed schedules more difficult to manage because workers have less scope to adjust work timing around fatigue and recovery needs. Consistent with evidence that health can condition workers' responses to flexible work arrangements [10], schedule-based flexibility may therefore be more strongly associated with subjective work performance among workers facing poorer health. This leads to the third hypothesis:

H3. *The positive association between schedule-based flexibility and subjective work performance is stronger among workers reporting poor health.*

The expected associations for location-based flexibility are less clear. Remote work may reduce commuting and increase autonomy, but it can also introduce communication, coordination, and boundary-management demands; moreover, policy availability and actual usage represent distinct stages of flexible-work implementation. Because these competing mechanisms and the existing evidence do not provide a clear basis for a directional prediction, remote-work availability and reported usage are examined exploratorily, both in comparison with schedule-based flexibility and across workers' health-status groups.

4. Data and Empirical Strategy

4.1. Data and Sample

This study uses data from the Japan Household Panel Survey and the Keio Household Panel Survey (JHPS/KHPS), nationally representative panel surveys covering employment, household characteristics, health status, work arrangements, and subjective evaluations of work and life. The panel structure provides repeated observations of the same individuals over time, allowing the analysis to examine associations between workplace flexibility, health status, and subjective work performance.

The baseline analysis uses the 2019–2021 annual survey waves, beginning in 2019, when the subjective work performance (SWP) measure was first introduced in the JHPS/KHPS. Because the JHPS/KHPS main survey is conducted in January, the 2019 and 2020 waves largely reflect pre-pandemic working conditions, whereas the 2021 wave captures an early pandemic-era setting. The baseline period thus covers the transition into the early stage of remote-work expansion, when location-based flexibility began to spread alongside the more established form of schedule-based flexibility (SBF). To assess whether the baseline associations carry over to later workplace conditions, the main SBF specification is also estimated separately for 2022–2024 and for the pooled 2019–2024 period (Section 5.5; Table A4). This comparison is relevant because evidence from the pandemic transition continues to inform post-pandemic debates on flexible work.

The sample is restricted to respondents aged 20–64 whose main activity in the previous month was work. The baseline regression sample further requires non-missing information on subjective work performance, self-rated health, and the controls included in the baseline specification. Workplace-flexibility indicators are constructed from the underlying survey responses using the affirmative-report coding described in Section 4.4. The resulting data form an unbalanced individual-level panel, with respondents appearing in one to three survey waves depending on survey participation and item availability.

Table 1 reports the sample construction and descriptive statistics for the main analytical sample. The baseline random-effects ordered logit models are estimated using 7678 person-year observations from 3259 individuals.

Table 1. Sample Construction and Descriptive Statistics.

Panel A. Sample Construction

Step	Restriction	Observations
1	Survey years 2019–2021	16,452
2	Age 20–64	11,114
3	Main activity was work last month	7983
4	Non-missing Subjective Work Performance and poor-health indicator	7810
5	Non-missing baseline controls	7678

Panel B. Descriptive Statistics for the Main Analytical Sample

Label	Mean	SD	Min	Max
Subjective Work Performance	6.16	1.73	0	10
Schedule-Based Flexibility (SBF)	0.18	0.39	0	1
Remote Work Exposure (RWE)	0.16	0.36	0	1
Remote-Work Availability (RWA)	0.13	0.33	0	1
Remote-Work Usage (RWU)	0.03	0.16	0	1
Poor Health	0.13	0.33	0	1
Age	45.65	10.92	20	64
Married (=1)	0.66	0.47	0	1
Male (=1)	0.62	0.49	0	1
Has Child(ren) (=1)	0.66	0.48	0	1
Wage/Salary employee (=1)	0.88	0.33	0	1
Observations	7678			

Panel C. Means by Poor-Health Status

Label	Mean, Poor health = 0	Mean, Poor health = 1	Difference
Subjective Work Performance	6.20	5.89	−0.31
Schedule-based Flexibility (SBF): 1	0.19	0.17	−0.02
Remote Work Exposure (RWE): 1	0.16	0.13	−0.03
Remote-Work Availability (RWA): 1	0.13	0.11	−0.03
Remote-Work Usage (RWU): 1	0.03	0.02	0.00
Age	45.29	48.11	2.82
Married (=1)	0.66	0.65	−0.01
Male (=1)	0.62	0.64	0.02
Has child(ren) (=1)	0.65	0.67	0.02
Wage/salary employee (=1)	0.88	0.88	0.00
Observations	6690	988	

Panel D. Person-Year Counts by Flexibility and Poor-Health Status

Flexibility Variable	Flexibility = 0, Poor Health = 0	Flexibility = 0, Poor Health = 1	Flexibility = 1, Poor Health = 0	Flexibility = 1, Poor Health = 1	Total
Schedule-based Flexibility (SBF): 1	5448	822	1242	166	7678
Remote Work Exposure (RWE): 1	5620	860	1070	128	7678
Remote-Work Availability (RWA): 1	5807	884	883	104	7678
Remote-Work Usage (RWU): 1	6503	964	187	24	7678

Notes: The baseline analytical sample consists of respondents aged 20–64 whose main activity in the previous month was work during 2019–2021. The final sample requires non-missing values for subjective work performance, poor health, and the baseline controls. The four flexibility indicators are constructed using the affirmative-report coding described in Section 4.4. SBF denotes Schedule-Based Flexibility; RWE denotes Remote Work Exposure; RWA denotes policy-only Remote-Work Availability; and RWU denotes Remote-Work Usage. Poor health equals one for respondents reporting self-rated health in the bottom two categories of the five-point scale; missing health responses are excluded. In Panel C, differences are calculated from unrounded means as Poor health = 1 minus Poor health = 0. Panel D reports person-year observations.

The analysis relies on the annual JHPS/KHPS main survey. The JHPS COVID-19 Special Survey, conducted from 2020 to 2022, is not combined with the main survey because its measurement design and sample composition differ.

4.2. Subjective Work Performance

The dependent variable is SWP. Respondents evaluate their overall work performance during the previous four weeks on an eleven-point scale from 0 to 10, with higher values indicating higher self-assessed work performance. This item was introduced in the JHPS/KHPS from the 2019 wave, which determines the earliest observation year in the present analysis. Accordingly, SWP is interpreted as respondents' subjective assessment of their recent work performance rather than as an objective measure of productivity.

Subjective work performance is particularly suitable for the present analysis because the JHPS/KHPS asks respondents to evaluate their overall work performance during the previous four weeks, providing a relatively proximal measure of recent work functioning. By contrast, available monetary labor-market outcomes such as annual earnings are more distal measures and reflect not only work performance but also working hours, job characteristics, contractual arrangements, and wage-setting practices. The SWP measure therefore more closely matches the present study's interest in short-run variation in work functioning associated with health status and workplace flexibility, although it should not be interpreted as an objective measure of productivity.

The main models treat subjective work performance as an ordinal outcome and are estimated using random-effects ordered logit models. For substantive interpretation, the estimates are also translated into predicted probabilities and average marginal effects.

4.3. Health Status

The key moderating variable is poor health, constructed from respondents' self-rated health. The indicator equals one when respondents report self-rated health in the bottom two categories of the five-point scale and zero when they report one of the upper three categories. Observations with missing self-rated health are excluded from the analytical sample. Self-rated health is used as a broad summary measure of respondents' overall health status relevant to work functioning.

The analysis treats poor health as a boundary condition that may shape the association between workplace flexibility and subjective work performance. Workers reporting poor health may face greater difficulty maintaining performance under less flexible work arrangements because health-related constraints can make fixed work schedules and commuting demands more difficult to accommodate. Accordingly, the association between workplace flexibility and subjective work performance may differ by health status.

4.4. Workplace Flexibility Measures

The analysis distinguishes four measures of workplace flexibility, capturing flexibility in work timing and work location as well as the distinction between remote-work availability and usage:

1. Schedule-Based Flexibility (SBF): Captures whether respondents report a working-time arrangement that permits flexibility or discretion over working hours. Based on the underlying working-time-system question, SBF equals one for respondents reporting flextime, discretion-based work, or a working arrangement without fixed time management, and zero otherwise. SBF therefore captures schedule-based flexibility embedded in the reported working-time arrangement rather than the degree of day-to-day schedule control actually exercised by the worker.

2. Remote Work Exposure (RWE): A broad indicator equal to one if the respondent reports either remote-work availability without usage or actual remote-work usage.
3. Remote Work Availability (RWA): Captures reported access to a remote-work arrangement without reported usage. The indicator equals one for respondents who report that remote work is available but do not report having used it.
4. Remote Work Usage (RWU): Captures reported actual use of remote work. The indicator equals one for respondents who report having used a remote-work arrangement.

The baseline coding follows an affirmative-report approach. Each flexibility indicator equals one when the respondent affirmatively reports the corresponding arrangement or usage and zero otherwise. The zero category should therefore be interpreted as not affirmatively reporting the relevant form of flexibility, rather than as necessarily confirming its absence. For the remote-work measures, "don't know" and missing responses are coded as zero in the baseline RWE, RWA, and RWU indicators; missing responses to the underlying working-time-system question are similarly coded as zero for SBF. Alternative treatments of these responses are examined in the robustness analysis.

Distinguishing these measures is central to the empirical design. SBF captures flexibility in work timing, whereas RWA and RWU capture different aspects of flexibility in work location. In addition, reported availability does not necessarily translate into actual use, and actual usage may be selective, reflecting job tasks, managerial expectations, workplace norms, household needs, and workers' preferences. The analysis therefore estimates separate models for each flexibility measure and compares their patterns of association with subjective work performance.

Because the baseline affirmative-report coding places unknown and missing responses in the zero category, robustness checks consider two alternative coding strategies. First, observations with missing responses to the underlying SBF question, or with "don't know" or missing responses to the underlying remote-work question, are excluded. Second, these responses are retained but modeled as separate response categories rather than included in the zero category. The results are reported in Table A6.

RWE is constructed from RWA and RWU and is therefore not entered jointly with those two components. Supplementary horse-race specifications instead compare SBF with RWE and, separately, SBF with RWA and RWU. These specifications assess whether the association between schedule-based flexibility and subjective work performance remains when location-based flexibility is also taken into account. The results are reported in Table A5.

4.5. Control Variables

The baseline models include a parsimonious set of individual, family, and job-related controls: gender, marital status, age and age squared, the presence of children, organization type, industry, firm size, employee status, job position, and educational attainment. Educational attainment is entered as a categorical variable, and all baseline models include year indicators.

The baseline specification is intentionally parsimonious to limit sample loss and avoid conditioning extensively on variables that may be closely related to the channels through which workplace flexibility is associated with subjective work performance. As a robustness check, a richer specification additionally controls for housework, childcare, and elder-care responsibilities; recurrent education; commuting frequency and commuting time; regional block and municipality size; and satisfaction with housing, safety, transport and shopping convenience, and the natural environment. This specification assesses whether the main SBF results are sensitive to more extensive covariate adjustment. The results are reported in Table A5.

4.6. Empirical Strategy

Because subjective work performance is measured on an ordinal scale from 0 to 10, the baseline specification uses a random-effects ordered logit model. This specification preserves the ordinal nature of the outcome while accommodating repeated observations in the short, unbalanced panel and using both within-person and between-person variation. The model assumes that the individual-specific random effect is normally distributed and uncorrelated with the included covariates. Because this assumption may be restrictive, correlated random-effects and linear individual fixed-effects specifications are examined as sensitivity checks (Section 4.7; Table A1). Standard errors are clustered at the individual level.

The baseline model is specified as follows:

$$y^*_{it} = \beta_1 F_{it} + \beta_2 H_{it} + \beta_3 (F_{it} \times H_{it}) + \mathbf{X_{it}}' \boldsymbol{\gamma} + \delta_t + a_i + \varepsilon_{it},$$

where y^*_{it} denotes the latent propensity underlying subjective work performance for individual i in year t. F_{it} denotes one of the four workplace flexibility measures defined in Section 4.4: SBF, RWE, RWA, or RWU. H_{it} is the poor-health indicator. $\mathbf{X}_{it}$ is a vector of control variables described in Section 4.5, δ_t denotes year indicators, a_i is the individual-specific random effect, and ε_{it} is the idiosyncratic error term.

The observed ordinal outcome is related to the latent propensity through estimated cut points

$$SWP_{it} = k \text{ if } \kappa_{k-1} < y^*_{it} \leq \kappa_k,\ k = 0, \ldots, 10,$$

where the lower and upper boundary cut points are defined as $\kappa_{-1} = -\infty$ and $\kappa_{10} = +\infty$, respectively. The coefficient β_1 captures the association between the flexibility measure and the latent propensity for higher subjective work performance among workers not reporting poor health. The coefficient β_2 captures the association between poor health and the latent propensity among workers without the relevant flexibility measure. The interaction coefficient β_3 indicates whether the association between flexibility and the latent propensity differs by health status.

Because an interaction coefficient in a nonlinear model does not directly reveal heterogeneity on the predicted-probability scale, the analysis reports predicted probabilities and average marginal effects (AMEs) in addition to the ordered-logit interaction coefficient [31]. The main probability-based outcome is the probability of reporting high subjective work performance, defined as SWP ≥ 7. Predicted probabilities are calculated by integrating over the estimated distribution of the individual-specific random effect. For the SBF model, average predicted probabilities are calculated for each combination of SBF and poor-health status, standardized over the same empirical covariate distribution of the baseline estimation sample. AMEs of SBF are then calculated separately by health status as discrete changes from SBF = 0 to SBF = 1, and the difference between the two health-status-specific AMEs is tested using the delta method.

The empirical strategy is designed to estimate panel-based conditional associations rather than causal effects. Workplace flexibility is not randomly assigned, and workers may select into flexible arrangements on the basis of job characteristics, health status, preferences, or workplace environments. The estimates should therefore be interpreted as associational evidence on how different forms of workplace flexibility are related to subjective work performance and whether these associations vary by health status.

4.7. Robustness and Sensitivity Checks

Several robustness and sensitivity checks are conducted to assess whether the baseline SBF findings depend on model assumptions, health measurement, missing data, the observation period, and alternative specifications. First, to assess sensitivity to unobserved time-invariant individual heterogeneity, the baseline random-effects ordered logit model

is compared with correlated random-effects (Mundlak) specifications [32] and a linear individual fixed-effects model. These specifications provide complementary checks on the baseline associations, although they do not eliminate concerns about time-varying confounding. The results are reported in Table A1.

Second, the sensitivity of the SBF × poor-health interaction to the measurement of health status is examined using alternative operationalizations of self-rated health. These include the original five-category health score, which permits a test of a linear health gradient in the interaction, and a three-category grouping that allows the SBF association to vary across health groups without imposing a linear gradient. The results are reported in Table A2.

Third, sample selection and missing-data sensitivity are examined through the sample-construction process, comparisons between included and excluded observations, and an alternative specification that retains missing baseline-control values as separate categories. Missing values for subjective work performance and health status are not imputed. These diagnostics and estimates are reported in Table A3.

Fourth, temporal heterogeneity is assessed by estimating the main SBF specification separately for 2019–2021 and 2022–2024, as well as for the pooled 2019–2024 period. A pooled model with interactions between SBF, poor health, and the survey period is used to formally test whether health-related heterogeneity differs between periods, including on the predicted-probability scale. Additional checks restrict the comparison to individuals observed in both periods and examine changes in sample composition and reported workplace arrangements. The results are reported in Table A4.

Fifth, the SBF specification is re-estimated with the richer set of controls described in Section 4.5 and with a sample restricted to wage and salary employees. Supplementary specifications also jointly include SBF and RWE or, alternatively, SBF, RWA, and RWU, to assess whether the SBF association persists when location-based flexibility is taken into account. RWE is not entered jointly with RWA and RWU because it is constructed from those components. The results are reported in Table A5.

Finally, alternative treatments of unknown and missing responses to the underlying flexibility questions are examined by excluding these responses or modeling them as separate response categories rather than incorporating them into the baseline zero category. The probability-based analysis is also repeated using a stricter definition of high subjective work performance (SWP ≥ 8), with corresponding predicted probabilities and health-status-specific average marginal effects. The results are reported in Table A6.

4.8. Use of Generative AI Tools

Generative AI tools were used to assist with manuscript drafting and language editing, suggestions for Stata coding and debugging, and checks of the clarity and consistency of the presentation and interpretation of statistical results. Prompts generally involved requests to revise author-provided text, suggest or debug statistical commands, and check descriptions of regression results against author-provided outputs. All statistical analyses were conducted and verified by the author. AI-generated suggestions were critically reviewed and edited as appropriate. Generative AI tools were not used to generate or alter the underlying research data or statistical results.

5. Results

5.1. Sample and Descriptive Patterns

Returning to Table 1, Panels B–D summarize the main descriptive patterns in the baseline analytical sample. Panel B reports descriptive statistics for subjective work performance, workplace flexibility, health status, and selected individual characteristics. Panel

C compares these characteristics by poor-health status, while Panel D reports person-year counts for each combination of flexibility status and poor-health status.

The mean subjective work performance score is 6.16 on the 0–10 scale, and approximately 13% of observations are classified as reporting poor health. Workers reporting poor health have a lower mean SWP score than those not reporting poor health (5.89 versus 6.20). The prevalence of the four flexibility measures also differs substantially: SBF is reported in approximately 18% of observations, compared with 16% for RWE, 13% for RWA, and 3% for RWU. These differences highlight that schedule-based flexibility and remote-work arrangements represent empirically distinct forms of workplace flexibility, and that remote-work availability is more common than reported usage.

Panel D further shows that actual remote-work usage among workers reporting poor health is rare in the 2019–2021 baseline sample. Only 24 person-year observations fall into the RWU = 1 and poor-health = 1 cell. Estimates involving RWU, particularly its interaction with poor health, are therefore likely to be imprecise and should be interpreted cautiously.

5.2. Main Ordered-Logit Results

Table 2 reports the baseline random-effects ordered-logit estimates. Each specification includes one workplace-flexibility measure and its interaction with poor health: Column (1) reports SBF, Column (2) RWE, Column (3) RWA, and Column (4) RWU. All specifications include the baseline controls and year indicators, with standard errors clustered at the individual level.

Table 2. Main Random-Effects Ordered-Logit Results.

	(1)	(2)	(3)	(4)
	SBF	**RWE**	**RWA**	**RWU**
Schedule-based flexibility	0.151 * (0.086)			
Remote-work exposure		−0.022 (0.086)		
Remote-work availability			−0.043 (0.088)	
Remote-work usage				0.067 (0.172)
Poor health	−0.407 *** (0.101)	−0.322 *** (0.100)	−0.304 *** (0.098)	−0.340 *** (0.093)
SBF × Poor health	0.472 ** (0.222)			
RWE × Poor health		−0.030 (0.233)		
RWA × Poor health			−0.195 (0.243)	
RWU × Poor health				0.607 (0.439)
Observations	7678	7678	7678	7678
Individuals	3259	3259	3259	3259
Log likelihood	−13,735.289	−13,740.698	−13,740.155	−13,739.595
Wald chi2	210.097	202.897	204.005	206.690

Notes: The dependent variable is subjective work performance, measured on an eleven-point scale from 0 to 10. The main analytical sample consists of respondents aged 20–64 whose main activity in the previous month was work during 2019–2021. All specifications include the baseline controls and year indicators. The baseline controls comprise gender, marital status, age, age squared, the presence of children, organization type, industry, firm size, wage/salary employee status, job position, and educational attainment. Educational attainment is entered as a categorical variable. Standard errors are clustered at the individual level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The first result is a consistent negative association between poor health and subjective work performance. Across all four specifications, the coefficient on poor health is negative and statistically significant. In the SBF specification, the coefficient is −0.407 (SE = 0.101, $p < 0.001$), indicating a lower latent propensity for higher subjective work performance among workers reporting poor health when SBF = 0. The corresponding poor-health coefficients are also negative and statistically significant in the remote-work specifications. This pattern is consistent with H1.

The second result concerns schedule-based flexibility. In the 2019–2021 baseline sample, the coefficient on SBF in Column (1) is positive and marginally significant (β = 0.151, SE = 0.086, $p = 0.080$). Because the model includes an interaction with poor health, this coefficient represents the association for workers not reporting poor health. The interaction between SBF and poor health is positive and statistically significant (β = 0.472, SE = 0.222, $p = 0.033$), indicating a stronger positive association between SBF and the latent propensity for higher subjective work performance among workers reporting poor health. The substantive magnitude of this health-related heterogeneity is examined on the predicted-probability scale in Section 5.3.

By contrast, the remote-work measures do not show similarly clear patterns in the 2019–2021 baseline sample. The coefficients on RWE and RWA are small and statistically insignificant, and their interactions with poor health are also statistically insignificant. The coefficient on RWU and its interaction with poor health are positive, but both estimates are imprecise and statistically insignificant. Given the very small RWU × poor-health cell noted in Section 5.1, these estimates should be interpreted cautiously. Overall, the baseline results indicate that SBF shows a clearer positive pattern in relation to subjective work performance than the remote-work measures.

5.3. Predicted Probabilities and Average Marginal Effects for Schedule-Based Flexibility

Because ordered-logit coefficients are not directly interpretable on the predicted-probability scale, Table 3 reports average predicted probabilities and average marginal effects (AMEs) for the baseline SBF model. The probability-based outcome is high subjective work performance, defined as SWP ≥ 7.

Table 3. Predicted Probabilities and Average Marginal Effects of Schedule-Based Flexibility.

Panel A. Predicted Probabilities of High Subjective Work Performance, Pr(SWP ≥ 7)						
SBF/Health Status	Predicted probability	SE		95% CI Lower		95% CI upper
SBF = 0, Poor health = 0	0.454	0.008		0.439		0.469
SBF = 0, Poor health = 1	0.388	0.016		0.356		0.419
SBF = 1, Poor health = 0	0.479	0.014		0.452		0.506
SBF = 1, Poor health = 1	0.490	0.034		0.424		0.556
Panel B. Average Marginal Effects of SBF on Pr(SWP ≥ 7), by Health Status						
Health Status	AME	SE	z	*p*-value	95% CI lower	95% CI upper
AME of SBF (0→1), Poor health = 0	0.025 *	0.014	1.747	0.081	−0.003	0.053
AME of SBF (0→1), Poor health = 1	0.102 ***	0.036	2.834	0.005	0.031	0.173

Table 3. *Cont.*

Panel C. Difference in Average Marginal Effects by Health Status			
Contrast	Estimate	SE	*p*-value
AME, Poor health = 1 − AME, Poor health = 0	0.077 **	0.037	0.035

Notes: High subjective work performance is defined as SWP ≥ 7. Predicted probabilities and average marginal effects are calculated from the baseline random-effects ordered logit model for SBF using 7678 person-year observations from 3259 individuals. Panel A reports average predicted probabilities for each combination of SBF and poor-health status. Panel B reports average marginal effects of SBF, calculated as discrete changes from SBF = 0 to SBF = 1, separately by health status. Panel C reports the difference in these average marginal effects between workers reporting poor health and those not reporting poor health. Standard errors are calculated using the delta method. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimates are associational and should not be interpreted as causal effects. Predictions integrate over the estimated random-effects distribution and are standardized over the common covariate distribution of the baseline estimation sample.

Panel A of Table 3 shows that, among workers without SBF, the predicted probability of high subjective work performance is 45.4% for workers not reporting poor health and 38.8% for those reporting poor health, a difference of approximately 6.6 percentage points. Among workers with SBF, the corresponding predicted probabilities are 47.9% and 49.0%, respectively. Thus, in the 2019–2021 baseline model, the health-related difference in predicted high performance observed when SBF is absent is not apparent when SBF is present.

Panel B reports the AMEs of SBF by health status. In the 2019–2021 baseline model, SBF is associated with a 2.5 percentage-point increase in the probability of high subjective work performance among workers not reporting poor health ($p = 0.081$), compared with a 10.2 percentage-point increase among workers reporting poor health ($p = 0.005$). Panel C shows that the difference between these AMEs is 7.7 percentage points ($p = 0.035$). Thus, on the predicted-probability scale, the baseline results indicate a stronger positive SBF association among workers reporting poor health.

Figure 1 visualizes the baseline probability estimates in Table 3. When SBF is absent, the predicted probability of high subjective work performance is lower among workers reporting poor health; when SBF is present, the predicted probabilities for the two health groups are similar. This graphical pattern mirrors the health-related heterogeneity in the baseline estimates.

5.4. Comparison with Remote-Work Measures

In the baseline ordered-logit models, RWE shows neither a statistically significant main association with subjective work performance nor a statistically significant interaction with poor health ($\beta = -0.022$ and -0.030, respectively). RWA yields a similarly weak pattern: both its main coefficient ($\beta = -0.043$) and its interaction with poor health ($\beta = -0.195$) are negative and statistically insignificant. Overall, the baseline estimates for RWE and RWA are small and imprecise and do not reveal a clear association with subjective work performance or clear health-related heterogeneity.

RWU shows a more positive coefficient pattern, but the estimates are imprecise. The RWU coefficient is positive ($\beta = 0.067$), as is its interaction with poor health ($\beta = 0.607$), but neither estimate is statistically significant. As shown in Table 1, only 24 person-year observations fall into the RWU = 1 and poor-health = 1 cell. The comparatively large positive interaction estimate should therefore be interpreted cautiously because health-related heterogeneity in RWU cannot be estimated precisely from this small subgroup.

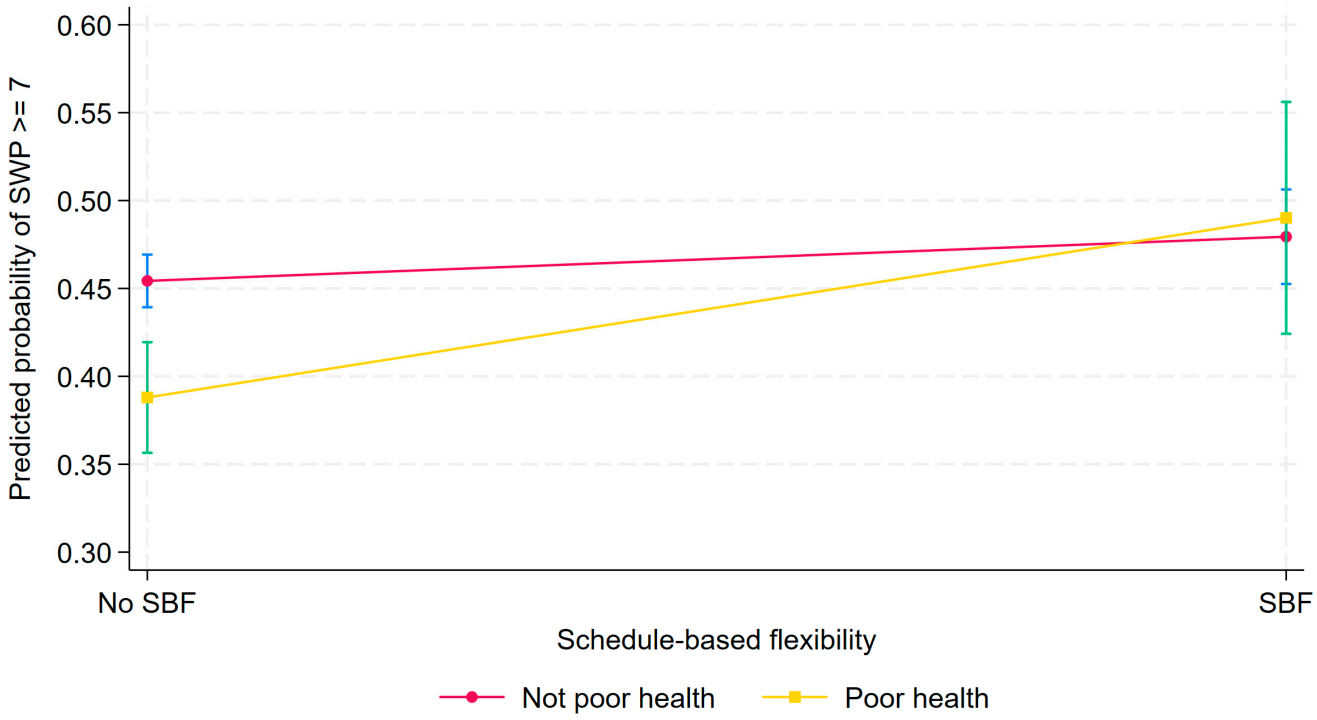


Figure 1. Predicted probability of high subjective work performance by schedule-based flexibility and health status. Notes: The figure presents average predicted probabilities of reporting high subjective work performance (SWP ≥ 7) by schedule-based flexibility and poor-health status. Predictions are calculated from the baseline random-effects ordered logit model using 7678 person-year observations from 3259 individuals. Blue and green vertical bars indicate 95% confidence intervals. The estimates are associational and should not be interpreted as causal effects.

Overall, the comparison across flexibility measures shows different empirical patterns across the forms of flexibility considered. In the 2019–2021 baseline analysis, SBF shows a clearer positive association with subjective work performance, while the remote-work estimates remain statistically imprecise, underscoring the value of distinguishing timing-based from location-based flexibility. At the same time, the statistically insignificant remote-work estimates should not be interpreted as evidence that remote work has no performance benefits. Rather, the present analysis does not provide sufficiently precise evidence of clear remote-work associations, particularly for health-related heterogeneity.

5.5. Results of Robustness and Sensitivity Checks

Tables A1–A6 assess the sensitivity of the baseline findings to unobserved individual heterogeneity, health measurement, missing data, temporal variation, model and sample specifications, and alternative response coding and performance thresholds. The main focus is whether the positive SBF × poor-health interaction observed in 2019–2021 persists under these alternatives.

Table A1 shows that the positive SBF × poor-health interaction in 2019–2021 persists under correlated random-effects specifications and a linear individual fixed-effects model, whereas the fixed-effects interaction for the pooled 2019–2024 period is small and statistically insignificant ($\beta = 0.047$, $p = 0.704$). The findings are more sensitive to the measurement of health (Table A2). The interaction with the five-category health score is not statistically significant ($p = 0.972$), whereas interactions in the three-category model are jointly significant ($p = 0.019$), indicating heterogeneity that is not well described by a linear health gradient. Table A3 documents differences between included and excluded observations in several

observed characteristics. Retaining missing baseline-control values as separate categories yields a positive but less precisely estimated interaction ($\beta = 0.430$, $p = 0.051$).

Temporal sensitivity is particularly consequential (Table A4). In separate-period models, the SBF × poor-health interaction is positive in 2019–2021 ($\beta = 0.472$, $p = 0.033$) but negative in 2022–2024 ($\beta = -0.623$, $p = 0.005$), whereas the pooled 2019–2024 interaction is small and statistically insignificant ($\beta = -0.062$, $p = 0.679$). A pooled model with period interactions confirms a significant difference between periods on the latent scale ($\beta = -0.808$, $p = 0.002$) and in the difference between health-status-specific SBF average marginal effects (−13.3 percentage points, $p = 0.002$). The reversal is also observed among the 2231 individuals represented in both periods. These results indicate substantial temporal heterogeneity in the conditional association, without identifying its causes.

Additional specifications provide qualified support for the baseline-period pattern (Table A5). The SBF × poor-health interaction remains positive with richer controls ($\beta = 0.431$, $p = 0.055$) and when RWE ($\beta = 0.486$, $p = 0.032$) or RWA and RWU ($\beta = 0.452$, $p = 0.051$) are included alongside SBF. It is smaller and statistically insignificant in the employee-only sample ($\beta = 0.256$, $p = 0.368$). Alternative treatments of unknown and missing flexibility responses leave the SBF interaction positive, while remote-work estimates generally remain imprecise (Table A6, Panels A–C). With a stricter definition of high performance (SWP ≥ 8), the SBF average marginal effect is 1.96 percentage points among workers not reporting poor health and 7.63 percentage points among those reporting poor health; the difference is 5.67 percentage points ($p = 0.0498$; Table A6, Panel D).

Taken together, the checks support a positive SBF × poor-health interaction in several 2019–2021 specifications, but they do not support interpreting this pattern as stable across health definitions, worker samples, or survey periods. These findings delimit the generalizability of the baseline result and motivate a cautious interpretation of the relationship between schedule-based flexibility, health status, and subjective work performance.

6. Discussion and Conclusions

6.1. Main Findings and Interpretation

This study examined whether different forms of workplace flexibility are associated with subjective work performance and whether these associations vary by workers' health status. Three findings emerge. First, poor self-rated health is consistently associated with lower subjective work performance. Second, in the 2019–2021 baseline period, SBF shows the clearest positive association among the flexibility measures considered, with a stronger association among workers reporting poor health. However, the SBF × poor-health interaction is negative in 2022–2024 and close to zero in the pooled 2019–2024 analysis, indicating substantial temporal heterogeneity. Third, the remote-work measures yield statistically imprecise associations, limiting conclusions about their relationship with work performance.

The baseline SBF findings are consistent with the possibility that schedule flexibility operates as a work-design resource. From a Job Demands–Resources perspective, greater discretion over work timing may help workers manage demands and maintain work functioning [8,9]. Such discretion may be especially relevant when health-related constraints affect fatigue, recovery needs, concentration, or the timing of work capacity. However, these potential mechanisms are not directly tested in this study. Moreover, the SBF × poor-health interaction reverses sign in 2022–2024, indicating that the baseline pattern should not be interpreted as evidence of a stable compensatory association across survey periods. This reversal may reflect changes in workplace arrangements, the implementation of schedule flexibility, or the composition of the workforce across survey periods. Although the reversal

is also observed among individuals represented in both periods, the present analysis cannot identify the mechanisms underlying this temporal variation.

In the 2019–2021 baseline analysis, the contrast between SBF and the remote-work measures highlights the distinction between timing-based and location-based flexibility. Remote work may reduce commuting burdens, but it does not necessarily provide discretion over when work is performed and may introduce coordination, communication, or work–nonwork boundary-management demands [6,22,25]. However, the remote-work estimates are generally imprecise, and these findings do not establish that location-based flexibility is ineffective or inferior to SBF. Rather, the results underscore the importance of examining distinct dimensions of workplace flexibility separately, rather than treating them as interchangeable work-design arrangements.

6.2. Implications for Sustainable Workplace Design

The findings have implications for sustainable workplace design, particularly its human and social dimensions. Sustainable work systems are not concerned only with short-term organizational performance, but also with whether work arrangements enable employees to maintain functioning and well-being over time [16,33]. From this perspective, the consistent negative association between poor self-rated health and subjective work performance highlights the importance of designing work arrangements that remain compatible with heterogeneous health-related constraints.

One implication is that schedule-based flexibility deserves attention alongside more visible forms of flexibility such as remote work. The present results do not establish that schedule-based flexibility causally improves performance, and the period-specific change in the SBF × poor-health interaction cautions against assuming a stable compensatory association. For workers whose energy, concentration, or recovery needs vary, greater flexibility over when work is performed may provide scope to adjust work rhythms without necessarily changing where work is performed. This is consistent with the broader view that work design should consider the resources employees can draw upon to manage job demands [8,9].

A second implication is that workplace flexibility should not be treated as a single organizational practice. Schedule flexibility and location flexibility provide control over different dimensions of work, and formal remote-work availability does not necessarily imply effective or usable flexibility [6,7]. Accordingly, remote-work policies should not automatically be regarded as substitutes for discretion over work scheduling. The relevant question is not simply whether flexibility exists, but whether the form of flexibility provided is usable under employees' actual work and health conditions.

These implications should remain modest. The study does not directly measure long-term employee retention, future health, or organizational resilience, and therefore does not demonstrate sustainability outcomes themselves. Rather, it suggests that SBF warrants further investigation as a potentially relevant component of health-supportive and inclusive work design.

6.3. Limitations and Future Research

Several limitations qualify the interpretation of the findings and suggest directions for future research. First, the analysis is observational and does not identify causal effects of workplace flexibility. Flexible work arrangements are not randomly assigned, and workers may differ in job characteristics, workplace environments, health status, preferences, or other factors related to subjective work performance. Although the baseline random-effects models account for repeated observations and observed covariates, and the correlated random-effects and linear fixed-effects specifications assess sensitivity to time-

invariant individual heterogeneity, unobserved selection and time-varying confounding may remain. Future research could strengthen identification by exploiting policy changes, organizational reforms, or other sources of plausibly exogenous variation in access to flexible work arrangements.

Second, the outcome is subjective work performance rather than objective productivity. This measure is useful for examining workers' own assessments of recent work functioning, including variation that may not be immediately reflected in wages, hours, or firm-level output. However, subjective assessments may also reflect reporting styles, expectations, morale, or other individual characteristics. Because SWP, self-rated health, and workplace-flexibility measures are all self-reported, common-method bias and measurement error may affect the estimated associations. The robustness checks address selected measurement choices but cannot fully rule out correlated reporting errors. The findings should therefore not be interpreted as direct evidence on objective productivity. Future studies could combine subjective performance measures with supervisor assessments, administrative performance indicators, absenteeism, or other objective work outcomes.

Third, the estimated health-related heterogeneity is sensitive to health measurement, sample restrictions, and the observation period. Although the positive SBF × poor-health interaction is observed in several 2019–2021 specifications, alternative health measures do not support a simple monotonic gradient, and the interaction reverses sign in the 2022–2024 period-specific model. The interaction is also estimated imprecisely in the employee-only sample. In addition, actual remote-work usage among workers reporting poor health is uncommon in the baseline sample, limiting the precision of the corresponding estimates. These patterns limit the generalizability of the baseline findings across health definitions, worker groups, and survey periods. The comparison between 2019–2021 and 2022–2024 documents temporal heterogeneity but does not identify the precise timing or causes of the observed changes in associations. Because 2024 is the most recent JHPS/KHPS wave currently available, the analysis cannot establish how these associations evolved during 2025–2026 or whether the period-specific patterns remain relevant to current workplace conditions. Future research using more recent survey waves, larger samples, and longer follow-up could examine whether these patterns recur under different labor-market and workplace conditions.

Finally, the flexibility measures capture reported workplace arrangements rather than the intensity or day-to-day exercise of flexibility. SBF identifies working-time arrangements that permit schedule flexibility but does not measure how much control workers actually exercise over their schedules. Similarly, the remote-work measures distinguish reported availability from usage but do not capture the frequency, intensity, or quality of remote work. The analysis also cannot identify the mechanisms through which schedule and location flexibility may be associated with work performance. Future research should therefore measure realized schedule control, remote-work intensity, managerial support, task characteristics, and related work-design features more directly, while examining mechanisms such as commuting burdens, recovery opportunities, autonomy, coordination costs, and work–nonwork boundaries.

6.4. Conclusions

This study provides evidence that different dimensions of workplace flexibility show different patterns of association with subjective work performance in Japan. Poor self-rated health is consistently associated with lower subjective work performance. In the 2019–2021 baseline analysis, SBF shows the clearest positive pattern among the flexibility measures considered, particularly among workers reporting poor health. However, the SBF × poor-health interaction reverses sign in 2022–2024 and is close to zero in the

pooled 2019–2024 analysis, indicating that this health-related pattern is not stable across survey periods.

These findings highlight the importance of distinguishing flexibility over when work is performed from flexibility over where work is performed. For sustainable workplace design, formal remote-work availability alone does not capture the full range of flexibility that may matter for workers' functioning. SBF may represent one relevant component of health-supportive and inclusive work design, but its association with work performance should not be assumed to be uniformly favorable across health groups or survey periods. Given the observational design, these findings should be interpreted as conditional associations rather than causal effects.

Funding: This research was funded by the Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant Number 25K23155.

Institutional Review Board Statement: Ethical review and approval were not required for this study because it involved secondary analysis of anonymized JHPS/KHPS microdata provided for academic research by the Panel Data Research Center, Institute for Economic Studies, Keio University. The author did not recruit participants, interact with survey respondents, or collect new personal data.

Informed Consent Statement: Not applicable to the present secondary analysis. The author did not recruit or interact directly with participants; the original survey procedures were administered by the Panel Data Research Center, Institute for Economic Studies, Keio University.

Data Availability Statement: Restrictions apply to the availability of these data. The JHPS/KHPS microdata used in this study were obtained from the Panel Data Research Center, Institute for Economic Studies, Keio University, and cannot be redistributed by the author under the applicable data-use conditions. Researchers may apply directly to the Panel Data Research Center for access to the JHPS/KHPS data through its Data Management System, subject to approval and the applicable terms of use. Information on data access is available at https://www.pdrc.keio.ac.jp/en/paneldata/howto/ (accessed on 28 September 2026).

Acknowledgments: The author gratefully acknowledges the helpful comments and suggestions received from participants at the Winter Workshop on Operations Research, Finance and Mathematics, 2026. The data for this analysis, the Japan Household Panel Survey and the Keio Household Panel Survey (JHPS/KHPS), were provided by the Panel Data Research Center, Institute for Economic Studies, Keio University. During the preparation of this manuscript, the author used ChatGPT (OpenAI; including GPT-5) for language editing and drafting assistance, coding and debugging suggestions, and assistance in checking the clarity and consistency of the interpretation and presentation of statistical results. The author reviewed and edited all outputs and takes full responsibility for the content of this publication.

Conflicts of Interest: The author declares no conflicts of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

JHPS/KHPS	Japan Household Panel Survey/Keio Household Panel Survey
SBF	Schedule-Based Flexibility
RWE	Remote Work Exposure
RWA	Remote-Work Availability
RWU	Remote-Work Usage
SWP	Subjective Work Performance
AME	Average Marginal Effect
JD–R	Job Demands–Resources

Appendix A

Table A1. Panel-Data Robustness to Unobserved Individual Heterogeneity.

Panel A. Random-Effects and Correlated Random-Effects Models			
Statistic	Baseline RE	Key-variable CRE	Full CRE
Schedule-Based Flexibility (SBF)	0.151 * (0.086)	0.007 (0.119)	0.029 (0.119)
Poor Health	−0.407 *** (0.101)	−0.233 * (0.138)	−0.247 * (0.139)
SBF × Poor Health	0.472 ** (0.222)	0.801 *** (0.255)	0.865 *** (0.252)
SBF AME, Non-poor health (pp)	2.513 * (1.438)	0.123 (1.963)	0.469 (1.956)
SBF AME, Poor Health (pp)	10.220 *** (3.607)	13.348 *** (4.258)	14.671 *** (4.185)
AME Difference: Poor—Non-poor (pp)	7.707 ** (3.657)	13.225 *** (4.169)	14.201 *** (4.100)
Person-Specific Means Included	No	SBF, poor health, interaction	All time-varying covariates
Joint Test of Included Person-Specific means	-	$p = 0.0035$	$p = 0.0334$
Observations	7678	7678	7678
Individuals	3259	3259	3259
Panel B. Within-Individual Fixed-Effects Robustness			
Statistic		2019–2021	2019–2024
SBF		0.019 (0.088)	0.063 (0.061)
Poor health		−0.158 (0.100)	−0.265 *** (0.066)
SBF × Poor health		0.558 *** (0.180)	0.047 (0.123)
SBF association, non-poor health		0.019 (0.088)	0.063 (0.061)
SBF association, poor health		0.577 *** (0.183)	0.109 (0.122)
Difference: poor—non-poor		0.558 *** (0.180)	0.047 (0.123)
Observations		7000	12,802
Individuals		2581	2793

Notes: Panel A reports random-effects ordered logit and correlated random-effects (Mundlak) specifications. AMEs refer to $\Pr(\text{SWP} \geq 7)$ and are reported in percentage points. Panel B reports linear individual fixed-effects estimates in SWP points; their magnitudes are not directly comparable with ordered-logit AMEs. Standard errors are reported in parentheses. Joint-test p-values are reported directly. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2. Sensitivity to Alternative Operationalizations of Self-Rated Health.

Panel A. Baseline Binary Poor-Health Specification	
Statistic	Estimate (SE)
Schedule-Based Flexibility (SBF)	0.151 * (0.086)
Poor health	−0.407 *** (0.101)
SBF × Poor health	0.472 ** (0.222)
SBF AME, non-poor health (pp)	2.513 * (1.438)
SBF AME, poor health (pp)	10.220 *** (3.607)
AME difference: poor—non-poor (pp)	7.707 ** (3.657)
Observations	7678
Individuals	3259

Table A2. *Cont.*

Panel B. Five-Category Self-rated Health Score	
Statistic	Estimate (SE)
SBF	0.193 (0.218)
Self-rated health score (higher = poorer health)	−0.273 *** (0.038)
SBF × Self-rated health score	0.003 (0.083)
SBF AME, health category 1 (pp)	3.291 (2.434)
SBF AME, health category 2 (pp)	3.346 ** (1.536)
SBF AME, health category 3 (pp)	3.355 ** (1.596)
SBF AME, health category 4 (pp)	3.317 (2.508)
SBF AME, health category 5 (pp)	3.234 (3.581)
Observations	7678
Individuals	3259
Panel C. Three-Category Self-Rated Health	
Statistic	Estimate (SE)
SBF (reference health group: categories 1–2)	0.256 ** (0.111)
Health: category 3 vs. categories 1–2	−0.286 *** (0.066)
Health: categories 4–5 vs. categories 1–2	−0.570 *** (0.109)
SBF × category 3	−0.250 * (0.147)
SBF × categories 4–5	0.347 (0.237)
Joint test of SBF × health interactions	$p = 0.0193$
SBF AME, health categories 1–2 (pp)	4.288 ** (1.868)
SBF AME, health category 3 (pp)	0.087 (1.878)
SBF AME, health categories 4–5 (pp)	9.930 *** (3.635)
AME difference: categories 4–5—categories 1–2 (pp)	5.642 (3.938)
Observations	7678
Individuals	3259

Notes: All models are random-effects ordered logit models with individual-clustered standard errors and the baseline controls. Higher values of the five-category self-rated health measure indicate poorer health. AMEs refer to Pr(SWP ≥ 7) and are reported in percentage points. The five-category interaction tests a linear health gradient; the three-category model allows the SBF association to vary flexibly across health groups. Joint-test *p*-values are reported directly. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3. Sample Selection and Missing-Data Diagnostics.

Panel A. Sample Construction and Complete-Case Funnel		
Stage	Observations	Individuals
2019–2021 observations	16,452	6208
Age 20–64	11,114	4303
Main activity in previous month = work	7983	3342
SWP and poor-health indicator observed	7810	3296
Final complete-case estimation sample	7678	3259
Excluded from working sample to final sample	305	-
Missing SWP and/or health	173	-
Additional observations with missing baseline controls	132	-

Table A3. *Cont.*

Panel B. Included versus Excluded Observations			
Characteristic	Excluded	Included	Difference
Age (years)	45.597	45.650	−0.053
Married (%)	56.1	66.0	−9.9 ***
Male (%)	59.3	62.1	−2.8
Has child(ren) (%)	63.9	65.6	−1.6
Schedule-Based Flexibility (%)	14.1	18.3	−4.2 *
Wage/salary employee (%)	81.9	87.9	−6.0 ***
Subjective Work Performance, where observed	5.731	6.158	−0.426 ***
Poor health (%), where observed	18.8	12.9	5.9 ***
Survey year 2019 (%)	51.1	37.6	13.6
Survey year 2020 (%)	26.6	33.7	−7.2
Survey year 2021 (%)	22.3	28.7	−6.4

Panel C. Complete-Case versus Missing-Control-Category Specification		
Statistic	Complete case	Missing-control categories
Schedule-Based Flexibility (SBF)	0.151 * (0.086)	0.159 * (0.086)
Poor health	−0.407 *** (0.101)	−0.378 *** (0.100)
SBF × Poor health	0.472 ** (0.222)	0.430 * (0.220)
SBF AME, non-poor health (pp)	2.513 * (1.438)	2.632 * (1.425)
SBF AME, poor health (pp)	10.220 *** (3.607)	9.636 *** (3.566)
AME difference: poor—non-poor (pp)	7.707 ** (3.657)	7.004 * (3.621)
Missing baseline-control values retained as categories	No	Yes
Observations	7678	7810
Individuals	3259	3296
Observations recovered	-	132

Notes: Panel A reports the construction of the 2019–2021 baseline sample. Of 7983 working person-year observations, 305 (3.82%) are excluded from the final complete-case model: 173 because SWP and/or poor-health status is missing and a further 132 because at least one baseline control is missing. Missing control values occur only for organization type, industry, wage/salary employee status, and education. Panel B reports descriptive, unadjusted comparisons; stars refer to tests of differences between excluded and included observations. The wage/salary employee, SWP, and poor-health rows use available observations. The differences in Panel B are unadjusted. In a diagnostic logit model for exclusion from the complete-case sample, controlling for age, sex, and SBF, the survey-year indicators are jointly significant ($p < 0.001$), whereas the SBF coefficient is not statistically significant ($p = 0.126$). Panel C retains missing baseline-control values as explicit categories; no SWP or health values are imputed. The recovered 132 observations have lower SBF prevalence, higher poor-health prevalence, and lower mean SWP than the complete-case sample. AMEs refer to $\Pr(\text{SWP} \geq 7)$ and are reported in percentage points. Standard errors are reported in parentheses in Panel C. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4. Temporal Heterogeneity in the SBF-Health Association.

Panel A. Separate-Period and Full-Period Models			
Statistic	2019–2021	2022–2024	2019–2024
Schedule-Based Flexibility (SBF)	0.151 * (0.086)	0.271 *** (0.101)	0.164 ** (0.067)
Poor health	−0.407 *** (0.101)	−0.612 *** (0.107)	−0.463 *** (0.074)
SBF × Poor health	0.472 ** (0.222)	−0.623 *** (0.224)	−0.062 (0.150)
SBF AME, non-poor health (pp)	2.513 * (1.438)	4.348 *** (1.614)	2.761 ** (1.135)
SBF AME, poor health (pp)	10.220 *** (3.607)	−5.265 * (3.106)	1.662 (2.383)
AME difference: poor—non-poor (pp)	7.707 ** (3.657)	−9.614 *** (3.337)	−1.099 (2.462)
Observations	7678	6342	14,020
Individuals	3259	2983	4011

Table A4. *Cont.*

Panel B. Formal Tests of Temporal Heterogeneity		
Statistic	Estimate (SE)	95% CI
SBF × Poor health, 2019–2021 (latent scale)	0.376 * (0.209)	[−0.034, 0.786]
SBF × Poor health, 2022–2024 (latent scale)	−0.432 ** (0.188)	[−0.800, −0.064]
Period difference in SBF × Poor health (latent scale)	−0.808 *** (0.265)	[−1.328, −0.288]
Pooled AME difference, 2019–2021 (pp)	6.231 * (3.505)	[−0.639, 13.101]
Pooled AME difference, 2022–2024 (pp)	−7.098 ** (2.999)	[−12.975, −1.220]
Formal period difference in AME contrasts (pp)	−13.329 *** (4.350)	[−21.856, −4.802]
Panel C. Common-Person Comparison		
Statistic	2019–2021	2022–2024
Schedule-Based Flexibility (SBF)	0.071 (0.101)	0.296 *** (0.110)
Poor health	−0.427 *** (0.117)	−0.614 *** (0.113)
SBF × Poor health	0.780 *** (0.250)	−0.593 ** (0.231)
SBF AME, non-poor health (pp)	1.166 (1.659)	4.748 *** (1.762)
SBF AME, poor health (pp)	13.869 *** (4.024)	−4.499 (3.237)
AME difference: poor—non-poor (pp)	12.703 *** (4.066)	−9.247 *** (3.498)
Observations	5956	5511
Individuals	2231	2231

Panel D. Sample Composition and Workplace Arrangements			
Characteristic	2019–2021	2022–2024	Difference
Schedule-Based Flexibility (%)	18.3	19.6	1.3
Remote Work Exposure (%)	15.6	31.2	15.6
Remote-Work Availability (%)	12.9	23.3	10.5
Remote-Work Usage (%)	2.7	7.8	5.1
Poor health (%)	12.9	14.3	1.5
Subjective Work Performance, mean	6.158	6.253	0.095
Age, mean	45.65	46.36	0.71
Married (%)	66.0	65.1	−0.9
Male (%)	62.1	60.1	−2.1
Has child(ren) (%)	65.6	65.0	−0.6
Wage/salary employee (%)	87.9	88.6	0.7

Notes: Panels A and C report random-effects ordered logit models with individual-clustered standard errors and the baseline controls. AMEs refer to $\Pr(\mathrm{SWP} \geq 7)$ and are reported in percentage points. Panel B reports formal temporal-heterogeneity tests from a pooled full-factorial SBF × poor-health × period model estimated over 2019–2024. Accordingly, estimates for both periods are jointly re-estimated using the extended sample. Pooled probability-scale margins are standardized separately within the observed covariate distribution of each period and therefore need not equal the separate-period AMEs exactly. Panel C restricts the analysis to 2231 individuals observed in both 2019–2021 and 2022–2024; 1028 individuals appear only in the earlier period, and 752 appear only in the later period. Panel D is descriptive; differences are calculated as 2022–2024 minus 2019–2021 using unrounded means. RWE rose sharply between the earlier and later survey years, whereas SBF changed more modestly. These period differences are associational and should not be interpreted as causal effects of the pandemic or of remote-work expansion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5. Additional Specification Robustness.

Panel A. Covariate and Sample-Definition Robustness			
Statistic	Baseline	Full controls	Employee-only
Schedule-Based Flexibility (SBF)	0.151 * (0.086)	0.152 * (0.086)	0.118 (0.100)
Poor health	−0.407 *** (0.101)	−0.280 *** (0.103)	−0.417 *** (0.106)
SBF × Poor health	0.472 ** (0.222)	0.431 * (0.224)	0.256 (0.285)
Observations	7678	7472	6749
Individuals	3259	3218	2894
Panel B. Horse-race Specifications of Workplace Flexibility			
Statistic		SBF + RWE	SBF + RWA + RWU
Schedule-Based Flexibility (SBF)		0.154 * (0.087)	0.150 * (0.088)
Poor health		−0.396 *** (0.105)	−0.390 *** (0.105)
SBF × Poor health		0.486 ** (0.227)	0.452 * (0.231)
Remote Work Exposure (RWE)		−0.036 (0.087)	-
RWE × Poor health		−0.099 (0.235)	-
Remote-Work Availability (RWA)		-	−0.044 (0.090)
RWA × Poor health		-	−0.199 (0.245)
Remote-Work Usage (RWU)		-	0.011 (0.179)
RWU × Poor health		-	0.375 (0.479)
Observations		7678	7678
Individuals		3259	3259

Notes: All specifications are random-effects ordered logit models with individual-clustered standard errors and year indicators. Panel A compares the locked 2019–2021 baseline specification with a richer-control specification and a sample restricted to wage and salary employees. The richer specification additionally adjusts for housework, childcare, and elder-care responsibilities; recurrent education; commuting frequency and commuting time; regional block and municipality size; and satisfaction with housing, safety, transport and shopping convenience, and the natural environment. The employee-only specification omits the employee-status indicator. Panel B jointly includes SBF with location-based flexibility measures. RWE is not entered jointly with RWA and RWU because RWE is constructed from those components. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6. Sensitivity to Response Coding and the Definition of High Subjective Work Performance.

Panel A. Schedule-Based Flexibility			
Statistic	Main coding	Valid only	Missing separate
SBF	0.151 * (0.086)	0.128 (0.089)	0.148 * (0.087)
Poor health	−0.407 *** (0.101)	−0.436 *** (0.101)	−0.436 *** (0.101)
SBF × Poor health	0.472 ** (0.222)	0.511 ** (0.222)	0.511 ** (0.222)
Observations	7678	7539	7678
Individuals	3259	3229	3259
Panel B. Remote Work Exposure			
Statistic	Main coding	Valid only	DK/Missing separate
RWE	−0.022 (0.086)	−0.027 (0.092)	−0.049 (0.088)
Poor health	−0.322 *** (0.100)	−0.381 *** (0.110)	−0.358 *** (0.108)
RWE × Poor health	−0.030 (0.233)	0.025 (0.245)	0.010 (0.238)
Observations	7678	6688	7678
Individuals	3259	3058	3259

Table A6. *Cont.*

Panel C. Joint SBF, RWA, and RWU			
Statistic	Main coding	Valid remote only	DK/missing separate
SBF	0.150 * (0.088)	0.154 (0.094)	0.155 * (0.088)
SBF × Poor health	0.452 * (0.231)	0.582 ** (0.250)	0.454 * (0.233)
RWA	−0.044 (0.090)	−0.050 (0.096)	−0.072 (0.092)
RWA × Poor health	−0.199 (0.245)	−0.155 (0.258)	−0.158 (0.251)
RWU	0.011 (0.179)	0.003 (0.183)	−0.018 (0.180)
RWU × Poor health	0.375 (0.479)	0.362 (0.507)	0.404 (0.481)
Observations	7678	6688	7678
Individuals	3259	3058	3259
Panel D. High-SWP Threshold Sensitivity			
AME of SBF, percentage points		Pr(SWP ≥ 7)	Pr(SWP ≥ 8)
Non-poor health		2.513 * (1.438)	1.960 * (1.136)
Poor health		10.220 *** (3.607)	7.633 *** (2.853)
Difference: poor—non-poor		7.707 ** (3.657)	5.673 ** (2.892)
Panel D (continued). Predicted Probability of SWP ≥ 8			
SBF and health status		Pr(SWP ≥ 8), percent (SE)	
SBF = 0, non-poor health		23.52 (0.65)	
SBF = 0, poor health		18.71 (1.14)	
SBF = 1, non-poor health		25.48 (1.12)	
SBF = 1, poor health		26.34 (2.73)	

Notes: All models are random-effects ordered logit models with individual-clustered standard errors, baseline controls, and year indicators. Main coding defines the binary flexibility measures through affirmative responses; non-affirmative, unknown, or missing responses are in the zero category. Valid-only models exclude missing SBF responses or unknown/missing remote-work responses. Separate-category models retain these responses using additional response indicators and interactions with poor health. Panels A–C show ordered-logit coefficients and standard errors. Panel D shows average marginal effects and predicted probabilities in percentage points, standardized over the common 2019–2021 baseline estimation sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

1. National Institute of Population and Social Security Research. *Population Projections for Japan (2023 Revision): 2021 to 2070*; National Institute of Population and Social Security Research: Tokyo, Japan, 2023.
2. Ministry of Health, Labour and Welfare. *Analysis of the Labour Economy 2025: Towards Sustainable Economic Growth Under Labour Supply Constraints*; Ministry of Health, Labour and Welfare: Tokyo, Japan, 2025.
3. Kuroda, S.; Yamamoto, I. Why Do People Overwork at the Risk of Impairing Mental Health? *J. Happiness Stud.* **2019**, *20*, 1519–1538. [CrossRef]
4. Suka, M.; Yamauchi, T.; Wada, K.; Yanagisawa, H. Questionnaire Surveys on Working During Treatment in Japanese Employees and Employers. *Sangyo Eiseigaku Zasshi* **2019**, *61*, 59–68. [CrossRef] [PubMed]
5. Ministry of Land, Infrastructure, Transport and Tourism. *FY2025 Telework Population Survey: Survey Results*; Ministry of Land, Infrastructure, Transport and Tourism: Tokyo, Japan, 2026.
6. Kossek, E.E.; Perrigino, M.B.; Lautsch, B.A. Work-Life Flexibility Policies from a Boundary Control and Implementation Perspective: A Review and Research Framework. *J. Manag.* **2023**, *49*, 2062–2108. [CrossRef]
7. Harrop, N.; Jiang, L.; Overall, N. A Meta-Analysis of Antecedents and Outcomes of Flexible Working Arrangements. *J. Organ. Behav.* **2026**, *47*, 208–236. [CrossRef]
8. Demerouti, E.; Bakker, A.B.; Nachreiner, F.; Schaufeli, W.B. The Job Demands-Resources Model of Burnout. *J. Appl. Psychol.* **2001**, *86*, 499–512. [CrossRef]
9. Bakker, A.B.; Demerouti, E. The Job Demands-Resources Model: State of the Art. *J. Manag. Psychol.* **2007**, *22*, 309–328. [CrossRef]
10. Rudolph, C.W.; Baltes, B.B. Age and Health Jointly Moderate the Influence of Flexible Work Arrangements on Work Engagement: Evidence from Two Empirical Studies. *J. Occup. Health Psychol.* **2017**, *22*, 40–58. [CrossRef] [PubMed]

11. Ducas, J.; Daneau, C.; Bouqartacha, S.; Lecours, A.; Abboud, J.; Marchand, A.-A.; Descarreaux, M. The Impact of Telework on Absenteeism, Presenteeism, and Return to Work among Workers with Health Conditions: A Scoping Review. *Front. Public Health* **2025**, *13*, 1655200. [CrossRef] [PubMed]
12. Morgeson, F.P.; Humphrey, S.E. The Work Design Questionnaire (WDQ): Developing and Validating a Comprehensive Measure for Assessing Job Design and the Nature of Work. *J. Appl. Psychol.* **2006**, *91*, 1321–1339. [CrossRef] [PubMed]
13. Parker, S.K.; Morgeson, F.P.; Johns, G. One Hundred Years of Work Design Research: Looking Back and Looking Forward. *J. Appl. Psychol.* **2017**, *102*, 403–420. [CrossRef] [PubMed]
14. Baltes, B.B.; Briggs, T.E.; Huff, J.W.; Wright, J.A.; Neuman, G.A. Flexible and Compressed Workweek Schedules: A Meta-Analysis of Their Effects on Work-Related Criteria. *J. Appl. Psychol.* **1999**, *84*, 496–513. [CrossRef]
15. Nijp, H.H.; Beckers, D.G.J.; Geurts, S.A.E.; Tucker, P.; Kompier, M.A.J. Systematic Review on the Association Between Employee Worktime Control and Work-Non-Work Balance, Health and Well-Being, and Job-Related Outcomes. *Scand. J. Work Environ. Health* **2012**, *38*, 299–313. [CrossRef] [PubMed]
16. Zink, K.J. Designing Sustainable Work Systems: The Need for a Systems Approach. *Appl. Ergon.* **2014**, *45*, 126–132. [CrossRef] [PubMed]
17. Currie, J.; Madrian, B.C. Chapter 50 Health, Health Insurance and the Labor Market. *Handb. Labor Econ.* **1999**, *3*, 3309–3416. [CrossRef]
18. Johns, G. Presenteeism in the Workplace: A Review and Research Agenda. *J. Organ. Behav.* **2010**, *31*, 519–542. [CrossRef]
19. Kessler, R.C.; Barber, C.; Beck, A.; Berglund, P.; Cleary, P.D.; McKenas, D.; Pronk, N.; Simon, G.; Stang, P.; Ustun, T.B.; et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J. Occup. Environ. Med.* **2003**, *45*, 156–174. [CrossRef] [PubMed]
20. Künn-Nelen, A. Does Commuting Affect Health? *Health Econ.* **2016**, *25*, 984–1004. [CrossRef] [PubMed]
21. Stutzer, A.; Frey, B.S. Stress That Doesn't Pay: The Commuting Paradox. *Scand. J. Econ.* **2008**, *110*, 339–366. [CrossRef]
22. Gajendran, R.S.; Harrison, D.A. The Good, the Bad, and the Unknown About Telecommuting: Meta-Analysis of Psychological Mediators and Individual Consequences. *J. Appl. Psychol.* **2007**, *92*, 1524–1541. [CrossRef] [PubMed]
23. Bloom, N.; Liang, J.; Roberts, J.; Ying, Z.J. Does Working from Home Work? Evidence from a Chinese Experiment. *Q. J. Econ.* **2015**, *130*, 165–218. [CrossRef]
24. Bloom, N.; Han, R.; Liang, J. Hybrid Working from Home Improves Retention Without Damaging Performance. *Nature* **2024**, *630*, 920–925. [CrossRef] [PubMed]
25. Yang, L.; Holtz, D.; Jaffe, S.; Suri, S.; Sinha, S.; Weston, J.; Joyce, C.; Shah, N.; Sherman, K.; Hecht, B.; et al. The Effects of Remote Work on Collaboration Among Information Workers. *Nat. Hum. Behav.* **2022**, *6*, 43–54. Correction in *Nat. Hum. Behav.* **2022**, *6*, 164. https://doi.org/10.1038/s41562-021-01228-z. [CrossRef] [PubMed]
26. Brynjolfsson, E.; Horton, J.; Ozimek, A.; Rock, D.; Sharma, G.; TuYe, H.-Y. *COVID-19 and Remote Work: An Early Look at US Data*; NBER Working Paper Series; National Bureau of Economic Research: Cambridge, MA, USA, 2020.
27. Dingel, J.I.; Neiman, B. How Many Jobs Can Be Done at Home? *J. Public Econ.* **2020**, *189*, 104235. [CrossRef] [PubMed]
28. Morikawa, M. Work-from-Home Productivity During the COVID-19 Pandemic: Evidence from Japan. *Econ. Inq.* **2022**, *60*, 508–527. [CrossRef] [PubMed]
29. Morikawa, M. Productivity Dynamics of Work from Home: Firm-Level Evidence from Japan. *J. Evol. Econ.* **2024**, *34*, 465–487. [CrossRef]
30. Morikawa, M. Productivity Dynamics of Remote Work During the COVID-19 Pandemic. *Ind. Relat. A J. Econ. Soc.* **2023**, *62*, 317–331. [CrossRef]
31. Greene, W. Testing Hypotheses About Interaction Terms in Nonlinear Models. *Econ. Lett.* **2010**, *107*, 291–296. [CrossRef]
32. Mundlak, Y. On the Pooling of Time Series and Cross Section Data. *Econometrica* **1978**, *46*, 69–85. [CrossRef]
33. Kobayashi, K.; Eweje, G.; Tappin, D. Employee Wellbeing and Human Sustainability: Perspectives of Managers in Large Japanese Corporations. *Bus. Strategy Environ.* **2018**, *27*, 801–810. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.